# EVALUATING ONE SCREEN/ ONE CONTROL MULTIFUNCTION DEVICES IN VEHICLES

**Peter C. Burns**
**Patricia L. Trbovich**
**Joanne L. Harbluk**
**Tara McCurdie**
Transport Canada
Canada
Paper Number 05-0339

## ABSTRACT

Multifunction in-vehicle information systems are becoming increasingly prevalent in cars. These systems typically use a centrally located display and a single control device to carry out a variety of operations including navigation, communications, entertainment, and climate control. Advantages of these systems include: conservation of dashboard space, improved styling, function integration and flexible configuration of functions. The aim of this research was to investigate potential disadvantages of these systems. Given the quantity and complexity of the information these systems provide and the attention required to operate these devices, there is concern that they may be overly difficult and distracting to use while driving. Two 2004 European luxury vehicles containing multifunctional information systems were used in this study. Both systems consisted of a center-mounted LCD screen and a console-mounted primary control knob. A combination of human factors assessment techniques were used to assess the systems: 1) expert evaluations: the Transportation Research Laboratory (TRL) Checklist and heuristic evaluations, 2) user testing and 3) the occlusion test. Six human factors experts performed the expert evaluations and 12 drivers participated in the user testing and occlusion testing. Results from the expert ratings provided a detailed account of problems. Specifically, the information display format in System A helped drivers maintain a correct representation of system status and provided immediate feedback. System B, in contrast, was less successful in terms of providing informative menu labels, appropriate feedback and navigation aids. The number of tasks successfully completed was assessed for the two systems. An average of 82% passed the performance goal in System A, and only an average of 38% in System B. Although these issues are important to the design of any consumer product, they are critical to the operation of in-vehicle systems as they could impair driver performance and increase crash risk.

## INTRODUCTION

In order to add more functionality into cars without increasing dashboard clutter, manufacturers are developing one control/one display information systems. These information systems integrate multiple functionalities, such as climate, navigation, entertainment and communication, and can be accessed via a single control. This space saving results in benefits such as more freedom for options and styling that enhances aesthetic and market appeal. Whereas the older style of interfaces generated distraction by requiring drivers to use multiple knobs and buttons, these newer systems may contribute to cognitive distraction by requiring drivers to remember what mode they are in. To access information that was once a button press away, drivers must now navigate through multiple hierarchical menu structures. Research is needed to evaluate how the use of these systems impacts driver safety.

At present, multifunctional systems are being introduced into cars without any standard criteria for their design. There are also no standard methods for assessing their ease of use and safety of operation while driving. To ensure that unsafe devices are not added to cars, appropriate assessment methods are required. The ultimate goal would be an assessment procedure that would eventually become the basis for an objective performance standard.

In the present study, we evaluated the effectiveness of both usability and safety assessment methods. The assessments were made using: (1) two version of expert evaluations, (2) user testing, and (3) occlusion testing.
We were specifically interested in the methods' sensitivity to different tasks and different interfaces.

## Expert Evaluations

Expert evaluations are performed by human factors and vehicle safety specialists who test how well a system meets a set of safety and usability design guidelines. These guidelines are based on requirements such as standards for physical sizing, location of controls, labelling and display of information, as well best practices for "look and feel" and functionality of interfaces. For each identified problem, experts give a severity rating to guide re-engineering priorities and provide solutions. Examples of expert evaluations include the Transportation Research Laboratory (TRL) checklist (Stevens et al., 1999) and Heuristic Evaluations (Nielsen, 1994).

**TRL Checklist** The TRL checklist is a structured evaluation tool in checklist format for assessing the safety related features of an in-vehicle information system. It was developed based on accepted existing codes of practice and emerging international standards and has much in common with the European Statement of Principles. It is a low cost assessment technique that only requires a pen and an in-vehicle information system. Following the checklist assessment, assessors complete a final report detailing both the good and bad features of the system's design. Systems recognised as having major safety concerns or numerous minor safety concerns are considered to be less safe than systems that are rated as having fewer or less serious safety concerns.

**Heuristic Evaluation** Heuristics, or "rules of thumb", are general principles used to guide design decisions. A Heuristic Evaluation (HE) consists of having evaluators examine a user interface, usually in the context of typical user tasks, to generate a list of problems and associated heuristic violations (Nielsen, 1994). The purpose of this method is to identify problems that could hinder the ease of use of the system. Nielsen's 1994 list of 10 heuristics provides the best developed set of user interface principles for use when critiquing a system. This set of principles is based on a principal components analysis of the usability problems found in a number of studies of various user interfaces. Nielsen suggests that 3 to 5 evaluators usually result in approximately 75% of the overall usability problems being discovered. Heuristic evaluations were first developed to evaluate website interfaces but have

been applied in many other domains such as in the evaluation of in-vehicle devices. Both types of expert evaluations (i.e., TRL checklist and HEs) are inexpensive and can be performed quickly and easily. As such, they offer a valuable front-end design evaluation tool for the automotive sector. Although expert evaluations highlight specific instances of problems, their usefulness lies in their ability to yield a high-level indication of weak aspects of an application that need further scrutiny. Expert evaluations are often combined with other assessment methods such as user testing. Specifically, once experts have identified types of problems, user testing can be performed on features that are most critical and relevant to tasks likely to be performed on these systems.

## User Testing

In User Testing evaluations, participants interact with an interface while being observed by an experimenter. Specifically, users are asked to perform a given task and speak aloud as they interact with the system. The experimenter notes the mistakes that the user makes as well as the "play by play" verbal feedback given by the user. Videotaping the session ensures that no important information is lost and also provides a compelling video record of the specific problems encountered by the user. In contrast to the qualitative and subjective expert evaluation methods discussed thus far in this paper, a user test is an objective performance measure that aims to test a product or system against a predetermined set of high-level usability goals such as efficiency, effectiveness, and satisfaction. Usability testing was applied in the present study to verify that problems indicated by the two expert evaluation methods result in actual problems for target users. The tasks chosen for the user testing were a set of common difficult in-vehicle tasks (i.e., set address and point of interest).

## Occlusion Testing

The probablility of a crash has been shown to increase as a function of increasing visual demands imposed by in-vehicle systems (Wierwille & Tijerina, 1998). Measuring "eyes-off-road" time by having people drive while interacting with an in-vehicle device can be dangerous and difficult. The occlusion method was developed as an indirect measure (i.e., no driving required) of visual demand of an in-vehicle task (ISO, 2004). Participants

perform in-vehicle tasks while wearing occlusion goggles that intermittently block their view of the in-vehicle device. The occlusion interval simulates drivers having to take their eyes off the display to look back at the road while still being able to manually operate the in-vehicle system. The vision interval is 1.5 seconds and the occlusion interval is 2.0 seconds. During the occlusion interval, the in-vehicle displays and controls are not visible but operation of the controls is still permitted. The occlusion testing technique differentiates in-vehicle tasks that require more or less sustained visual attention to complete a task successfully. The key measure of sustained visual attention is the Total Shutter Open Time (TSOT) which is calculated by multiplying the number of vision intervals (i.e., shutters open) needed to complete the task by the 1.5 seconds vision interval. Tasks that can be completed in a few brief glances (i.e., shorter TSOT) are considered to be less visually distracting than tasks that require a greater number of glances (i.e., longer TSOT). Presently, there are no agreed upon specific performance criteria although these issues are being examined in an ISO draft work item. The Japanese Automobile Manufacturers Association (JAMA), however, has recently provided guidelines recommending a maximum TSOT of 7.5 s when a system is bench tested using the occlusion method.

## METHODOLOGY- EXPERT EVALUATIONS

### Evaluators

Six usability experts, working in pairs, performed 3 evaluations using the Transportation Research Laboratory (TRL) checklist and 3 heuristic evaluations. Three experts had background and experience in automotive human factors. The remaining three experts had combined backgrounds and experience in cognitive psychology, human-computer interaction and systems engineering. Because the three evaluators with expertise in automotive human factors were familiar with both multifunctional devices, they were each paired with one of the other three evaluators. Specifically, the combinations of expertise were as follows: pair #1- automotive human factors/cognitive psychology, pair #2- automotive human factors/human-computer interaction and, pair #3- automotive human factors/systems engineering. The evaluator familiar with the systems was able to acquaint the other evaluator with the system and describe the typical task scenarios in which the interface is used. The

same pair of evaluators assessed System A and B separately.

### Apparatus

Two European luxury vehicles (model year 2004) containing multifunctional information systems (System A and System B) were used in the evaluations. Both multifunctional information systems consisted of a centre-mounted Liquid Crystal Display (LCD) screen and a console-mounted main control knob that worked as the system's primary control. Both vehicles were stationary during testing.

### Procedure

Each team of evaluators began the evaluation with an introduction to the system provided by the evaluator most familiar with the system. After the explanation of the nature and purpose of the functions included in the multifunctional systems, the team proceeded with their systematic evaluation of the interface.

### Materials

**TRL Checklist** The TRL checklist used in the present study was developed by the Transport Research Laboratory for the UK Department for Transport. Prior to commencing the evaluation of the multifunctional interfaces, evaluators read the comprehensive instructions and detailed guidelines contained in the user manual that accompanies the TRL checklist. This manual contains supportive information providing: (1) an explanation about the application of the checklist, (2) the rationale for the questions contained in the checklist, (3) a list of technical references and abbreviation, and (4) a glossary of terms. Evaluators completed the 3 separate parts of the TRL checklist: (1) assessment scenario, (2) in-depth assessment, and (3) assessment summary.

**Heuristic Evaluation** The checklist guiding the evaluation contained 10 heuristics (see Nielsen 1994 for a review) that have been shown to cover the majority of usability problems users might encounter. The list functions as a reminder to the evaluator of potential problem categories. An example of one such heuristic "navigation" refers to the presence or absence of suitable navigation tools, presented in appropriate places, and leading to application areas that are consistent with the users'

expectations. The evaluators worked through a set of typical tasks identifying problems and their associated heuristic violations as these occurred.The result is a list of problems and their corresponding severity. The process can be taken a step further in that solutions can be proposed.

## RESULTS- EXPERT EVALUATIONS

### TRL Checklist

Results from the TRL checklist provided a detailed account of potential problems. Specifically, experts predicted that the way information was displayed in system A would help drivers maintain a correct representation of system status and provide immediate feedback. Conversely, experts predicted that System B placed inadequate emphasis on issues such as use of informative labels, appropriate feedback and navigation aids. Although these issues are important to the design of any system, they are critical to the operation of in-vehicle systems as they could impair driver performance by increasing the demands on the driver.

The greatest difference between the two systems, based on how they scored on the TRL checklist, was that visual information presentation was better for System A than for System B. The larger number of menus and menu layers on System B increased its complexity relative to System A. Experts judged that System B's design would make it more difficult for users to see where they were in the system, how they got there, and how to get back to the starting point. Experts also rated System B as being more difficult to return to the start or escape from a dead end. This problem was due to the inconsistency in the return and escape options. In sum, experts concluded that it would be more difficult to navigate System B's interface than System A's interface. The TRL checklist states that systems that are more difficult to navigate will require more visual interaction time. This hypothesis was tested during the occlusion testing, the results of which are discussed below.

### Heuristic Evaluation

As shown in Table 1, the total number of problems identified for System A was 35 and the total number of problems identified for System B was 51. Some problems identified violated more than one heuristic resulting in the number of violations exceeded the number of problems.

Both systems had a large number of heuristic violations given that these heuristics cover fairly

basic requirements. From Table 1, we can see that there were more heuristic violations in System B than in System A which suggests that System B is less easy to use than System A.

**Table 1**
**Number of Violations by Heuristic and System**

| Heuristic | System A | System B |
|---|---|---|
| 1. Visibility of system status | 11 | 17 |
| 2. Match between system and the real world | 10 | 26 |
| 3. Recognition rather than recall | 3 | 6 |
| 4. Consistency and standards | 12 | 18 |
| 5. User control and freedom | 9 | 7 |
| 6. Flexibility and efficiency | 0 | 6 |
| 7. Aesthetics and minimalist design | 5 | 9 |
| 8. Error prevention | 6 | 6 |
| 9. Help users recognize, diagnose, and recover from errors | 0 | 0 |
| 10. Help and documentation | 3 | 4 |
| **Total** | **59** | **99** |

The number and nature of heuristic violations give a global overview of problems. This overview is regarded as a first step in usability evaluation in which areas of concern are identified to guide further, more detailed usability evaluations and to highlight specific issues to be exposed in subsequent user testing. For example, System B's interface appeared to suffer from a lack of match between system and the real world. This finding signals a need to review all words, symbols, actions and concepts to ensure that they are familiar to users (rather than system-specific engineering terms) and to test the effect of one or two instances of the problem on user performance. Furthermore, the design of the interface's navigation should reflect the order in which users will most likely perform tasks. System B's interface also appeared lack of standardization and consistency. This finding signals that users may be confused as to whether different words, icons and actions mean the same thing in different situations. It is prefreable to follow a conventional platform when designing an interface.

Experts predicted that terminology would be a problem for locating and interpreting information on System B due to the observation that headings and sub-headings were often difficult to understand and the information contained under many of these headings might not meet users' expectations. In this way, the heuristic evaluation serves as a guide for deeper subsequent probing to ensure identification and removal of all instances of a given problem type.

In sum, results from the Heuristic Evaluation were consistent with findings from the TRL checklist. In both cases experts found more usability and safety issues with System B's interface than System A's interface. Once experts identified the potential usability and safety problems perceived to exist in these systems, user testing was conducted to determine the degree to which the problems impede the typical user's' ability to complete specific tasks.

## METHODOLOGY- USER TESTING

### Participants

Twelve participants (11 males and 1 female) took part in the user testing. The participants ranged in age from 25 to 57 years with a mean age of 40. All were experienced drivers with normal or corrected to normal vision.

### Materials

The same vehicles and multifunctional devices used for expert evaluations were used for the user testing. Both vehicles were stationary at all times.

### Procedure

Participants sat in the drivers seat of the stationary vehicle. An experimenter seated in the front passenger seat administered the tasks to the participants. The experimenter seated in the back of the vehicle video- recorded the session. Participants were first familiarized with how the multifunctional information system functioned and given a few minutes to review the system. The goal was to assess how easy it is for drivers to locate and interpret specific information in the system. They performed four tasks which were developed based on the features that are most critical and relevant to tasks likely to be performed on these systems. Specifically, the experimenter asked participants to perform the following tasks:

- Task 1- Set address as destination: Participants were given an Ottawa address and asked to enter the street name and street number into the navigation system as the destination.
- Task 2 - Manually tune radio station and store it: Participants were given a specific radio frequency and asked to manually search for and select it.
- Task 3 - Set point of interest as destination: Participants were given a specific place of interest (e.g., restaurant, hotel) in other cities (i.e., different from Ottawa) and were asked to search for that place of interest within the navigation system and input it as the destination.
- Task 4 - Adjust audio setting: Participants were asked to adjust different "Treble/Bass" or "Balance/Fader" settings.

Each participant attempted the four tasks three times for a total of 12 tasks using each of the systems. Task order and system used was counterbalanced across participants. Participants were asked to speak out loud about their actions as they performed each task. Individual sessions lasted up to one hour.

### Measures

The number of tasks completed successfully was the usability metric applied to all tasks. For a task to be completed successfully, users had to complete it making a maximum of two errors. If users made more than two errors, or they were unable to find the information, it was considered a failure.

## RESULTS- USER TESTING

Of the 4 main tasks, an average of 82% of participants passed the performance goal in System A, and only an average of 38% in System B. The following table shows a summary of the results.

- 8/12 drivers using System A, and 6/12 drivers using System B were able to set an address as their destination point.
- 9/12 drivers using System A, and 6/12 drivers using System B were able to manually tune the radio station and store it.
- 10/12 drivers using System A, and 1/12 drivers using System B were able to set a point of interest as their destination.

- 12/12 driver using System A, and 5/12 drivers using System B were able to adjust the audio to a given setting.

These are indicated as percentages in Table 2. The results from the user testing support the experts prediction that System B was more difficult to use than System A. For system B, the major usability issue was that participants didn't know what menu or sub-menu labels to look under to find the desired information. This result is also consistent with the finding from the Heuristic Evaluation that terminology appeared to be a problem for locating and interpreting information on System B.

**Table 2.**
**Percentage Completion by Task and System**

| Tasks | System A | System B |
|---|---|---|
| Set Address as Destination | 67% | 50% |
| Manually Tune Radio | 75% | 50% |
| Set Point of Interest as Destination | 83% | 8% |
| Adjust Audio | 100% | 42% |
| *Average* | *82%* | *38%* |

To a large extent, the success of any information system, such as an Intelligent Vehicle Information System (IVIS), will depend on its usability or ability to be easily understood and conveniently employed by a user. Another important factor to consider when evaluating these devices is their safety performance. To assess whether these systems differ in the safety they provide to users, a user testing employing the occlusion procedure was performed. If results from the user testing suggest that System B is less safe than System A in terms of the amount of visual resources needed to perform the tasks, it will be more compelling for designers to take more care and effort to improve the systems.

## METHODOLOGY- OCCLUSION TESTING

### Participants

The same 12 participants that participated in the user testing took part in the occlusion testing.

### Apparatus and Tasks

Liquid crystal shuttering spectacles were used to intermittently block the participant's vision (Translucent Technologies Inc. Toronto). The goggles were programmed such that the vision interval, with shutter open, was 1.5 seconds (within the suggested maximum time tolerance for having eyes off the road; Zwahlen et al., 1988) and the occlusion interval, with shutter closed, was 2.0 seconds. The same vehicles and tasks used for expert evaluations and user testing were used for the occlusion testing.

### Procedure

During the experimental task trials, participants sat in the driver seat of the stationary vehicle. An experimenter seated in the front passenger seat administered the tasks to the participants. The experimenter seated in the back of the vehicle recorded task completion times. Sessions were conducted during daylight hours.
Participants were familiar with the multifunctional in-vehicle devices from the previous user testing session. They were given three occlusion warm up tasks involving the climate control system to familiarize themselves with the goggles and viewing conditions. Participants were then presented with the experimental conditions where they performed 12 tasks (3 repetitions of the 4 tasks) while wearing the occlusion goggles. The lenses on the goggles alternated from clear to opaque at intervals of 1.5 seconds and 2 seconds respectively until task completion. Participants also performed the 12 tasks with the lenses open. Performance was timed for each task. System order, task order and occlusion order were counterbalanced across participants.

Before each task began, the goggles remained open while participants viewed instructions printed on a flash card. Participants were asked to signal that they had read and understood the instructions by saying 'OK', and then were asked to complete the requested task to the best of their ability using the system. The task ended when the participant had completed the task or when five minutes had elapsed (whichever came first).

The dependent variable of interest was the Total Shutter Open Time (TSOT), the total time that vision is not occluded when using the occlusion procedure. TSOT is the sum of vision intervals required to complete a given task (ISO, 2004) and is a surrogate for total eyes-off-road time.

## RESULTS- OCCLUSION TESTING

A 4 x 2 analysis of variance (ANOVA), with repeated measures on both factors (i.e., 4 task type; set address vs. tune radio vs. set point of interest vs. adjust audio and 2 devices; System A vs. System B) was conducted to test for differences in mean Total Task Time in the unoccluded condition ($TTT_{unocc}$). A significant interaction of task type and device type was observed [$F(3,33) = 30.02$, $p < 0.001$]. Post hoc pairwise comparisons using Fisher's LSD test revealed that for the "Set address as destination" task, the mean total task time was significantly higher when participants used System B (mean = 78.26sec.) than when they used System A (mean = 54.66 sec.). Similarly, for the "Set point of interest" task, the mean total task time was significantly higher when participants used System B (mean = 95.42 sec.) than when they used System A (mean = 48.22 sec.) (see Figure 1). These results suggest that the components involved in the "Set address as destination" and the "Set point of interest" tasks may be unsafe and require further scrutiny.
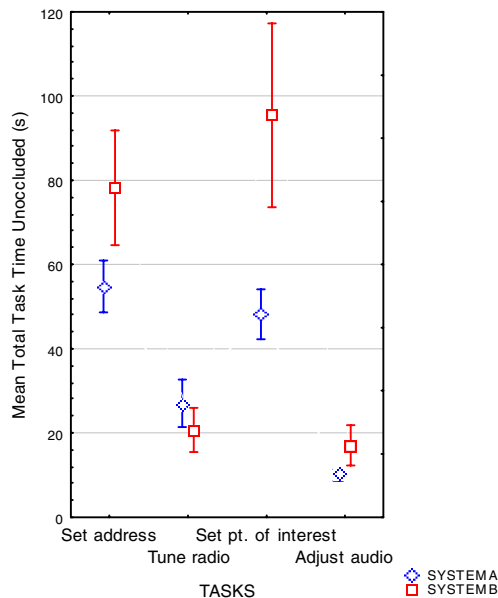


**Figure 1. Mean Total Task Time without Occlusion by Task and System.**

Results for the Total Shutter Open Time (TSOT) were similar to results for the total task time ($TTT_{unocc}$). Specifically, a 4 x 2 analysis of variance

(ANOVA) with repeated measures on both factors (i.e., 4 tasks; set address vs. tune radio vs. set point of interest vs. adjust audio and 2 devices; System A vs. System B) was conducted to test for differences in mean Total Shutter Open Time (TSOT). A significant interaction between task type and device type was observed [$F(3,33) = 34.15$, $p < 0.001$]. Post hoc pairwise comparisons using Fisher's LSD test revealed that for the "Set address as destination" task, the mean total shutter open time was significantly higher when participants used System B (mean = 58.67 sec) than when they used System A (mean = 46.79 sec). Similarly, for the "Set point of interest as destination" task, the mean total shutter open time was significantly higher when participants used System B (mean = 78.53 sec.) than when they used System A (mean = 35.93sec). These task times are quite long. To ensure safe operation of multifunctional information systems, complex operations such as setting an address or point of interest as a destination should be restricted by only being accessed when the vehicle is not in motion.
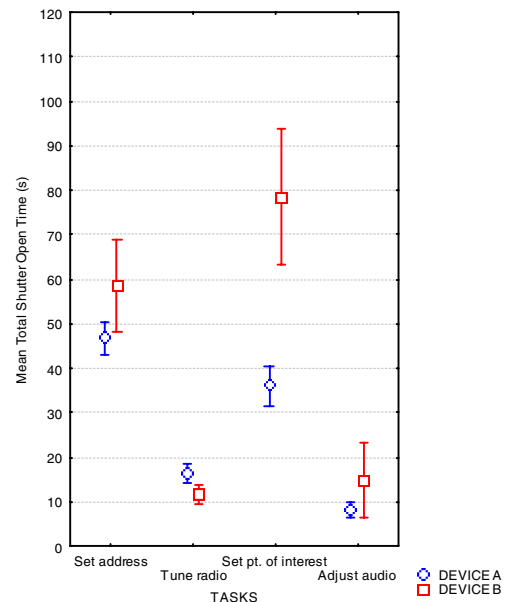


**Figure 2. Mean Total Shutter Open Time by Task and System.**

In sum, results for $TTT_{unocc}$ and TSOT show that the occlusion procedure was able to discriminate between the demands of the two different interfaces. Thus, the power of the occlusion procedure as a method for evaluating visual demands of in-car information systems is supported. These results also support the TRL statement that the system most difficult to navigate (i.e., System B) would also

require the most visual interaction time (TSOT System B sig. > TSOT System A). Further, a safety requirement for display-based in-vehicle systems is that the information must be quickly readable and understandable (Baumann et al. 2004). Results indicate that System A satisfied the latter requirement more so than System B, suggesting that System A is safer than System B. It is interesting to note that based on the JAMA guidelines for TSOT, none of these tasks would be considered safe or acceptable because they exceed their 7.5 seconds TSOT criteria.

## GENERAL DISCUSSION

Expert evaluations of the two multifunctional devices yielded a global overview of their associated problems and were valuable in identifying the number and nature of usability and safety violations. Specifically, System B showed more usability and safety violations than System A. This finding demonstrates the value of expert evaluations in discriminating the number of basic usability and safety problems between two multifunctional displays. The increased number of usability violations found in System B, relative to System A, was consistent with the subsequent user testing results which indicated that users had more difficulty performing tasks on System B than on system A. Thus, user testing contributed to the assessment process by validating assumptions from expert evaluations. Finally, the occlusion procedure proved to be a useful method for evaluating safety, by assessing the visual processing demands of the multifunctional displays. The results in terms of total task time ($TTT_{unocc}$) and total shutter open time (TSOT) clearly showed that System A was superior to System B for the two more complex tasks (i.e., "Set address as destination" and "Set point of interest" tasks).

The present findings provide an important perspective on the different roles of assessment methods in the evaluation of multifunctional in-vehicle interfaces. Expert evaluations and user testing of System A and System B accurately predicted superior safety performance of System A over System B. Given the latter and the fact that expert evaluations and user testing are cost effective and can be applied quickly, proper evaluation chronology should first conduct expert evaluations and user testing and then more defined tests such as

occlusion testing. Thus, to have the most impact on the usability of a system, expert evaluations and user testing should be incorporated into the early phases of the development process and continue as iterative testing during the remainder of the development process. Most developers acknowledge the value of usability testing, but many still view it as a hindrance to a timely and orderly product development process. The results from the present study suggest that these fears are justified when a usability evaluation serves only as a final checkpoint before the product is released to the public.

Given the number and seriousness of the problems found with the readily available systems evaluated in this study, one is lead to wonder why the developers did not catch these problems given that these techniques are simple and cost effective to implement.

Researchers have suggested that methods to evaluate safety and usability of multifunctional interfaces in cars are needed early on in the design process (Bullinger & Dangelmaier, 2003; Nowakowski et al., 2003). The results of the present research support this view and demonstrate that expert evaluations, user testing and occlusion testing provide a good combination of methods for assessing usability and safety of multifunctional information systems.

## CONCLUSION

While safety should be at the forefront of system design and evaluation, user requirements also need to be met. It is imperative that a balance is reached between safety and user requirements. There is a need to understand how drivers use functions and services provide by multifunctional systems. The input of human factors specialists early in the development would help ensure user requirements are examined and met so that IVIS s may even decrease driver workload if user needs are matched in a way that is compatible with the primary task of driving. Together, the expert evaluations (i.e., TRL checklist and heuristic evaluation), the user testing and the occlusion testing results can help designers identify the areas and seriousness of both usability and safety issues.

Although System A showed less usability and safety problems than System B, it is surprising and disappointing that both systems rated poorly on these safety and usability evaluations. There is clearly a need to incorporate usability and safety assessment

methodologies in the development of in-vehicle devices. If such methods are being used, then a better process is needed to place more importance on this information and to assure that problems are acknowledged, assigned and tracked until they have been resolved. Once evaluations become an integral part of the system development process, the end result is a safe and easy to use system.

More research is needed to validate and refine assessment methods. Specifically, assessment methods would benefit from criterion values for acceptable driver distraction. Thus, the next step will be to define some criteria on which to set performance limits for unsafe tasks.

## REFERENCES

[1] Baumann, M., A. Keinath., J. F. Krems, and K. Bengler. Evaluation of in-vehicle HMI using occlusion techniques: experimental results and practical implications. *Applied Ergonomics*, Vol. 35, 2004, pp. 197-205.

[2] Bullinger, H-J., and M. Dangelmaier. Virtual prototyping and testing of in-vehicle interfaces. *Ergonomics*, Vol. 46, 2003, pp. 41-51.

[3] ISO/TC22/SC13/WG8 NWI 16673. "Road vehicles—Ergonomics aspects of transport information and control systems —Occlusion Method to Assess Visual Distraction Due to the Use of In-Vehicle  Systems". Draft work item, 2004.

[4] JAMA 2004 Guideline for In-vehicle Display Systems- Version 3.0. Japanese Automobile Manufacturers Association.

[5] Nielsen, J. 1994. Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), "Usability Inspection Methods". John Wiley & Sons, New York, NY.

[6] Nowakowski, C., P.Green, and O. Tsimhoni. Common Automotive Navigation System Usability Problems and a Standard Test Protocol to Identify Them. Paper presented at the 13th Annual ITS America Meeting (CD-ROM), Washington, D.C. Intelligent Transportation Society of America, 2003.

[7] Stevens, A., Board, A., Allen, P., and Quimby, A. (1999). "A safety checklist for the assessment of in-vehicle information systems: a user manual"
Transport Research Laboratory. Published report number: PA3536/00.

[8] Wierwille, W. and Tijerina, L. (1998). Modelling the relationship between driver in-vehicle visual demands and accident occurrence.  In A.G. Gale, I.D. Brown, C.M. Haslegrave, and S.P. Taylor (Eds.) Vision in Vehicles VI (pp.233-244). Amsterdam: North-Holland.

[9] Zwahlen, H.T., C.C. Adams Jr., and D.P. DeBald. "Safety aspects of CRT touch panel controls in    automobiles". In: Gale, A.G., Freeman, M.H., Haslegrave, C.M., Smith, P., Taylor, S.P. (Eds.), Vision in Vehicles-II. Elsevier Science Publishers, 1988, pp.335-344.